Analysis of DGGE profiles to explore the relationship between prokaryotic community composition and biogeochemical processes in deep subseafloor sediments from the Peru Margin

John C. Fry[1], Gordon Webster[1,2], Barry A. Cragg[2], Andrew J. Weightman[1] & R. John Parkes[2]

[1]Cardiff School of Biosciences, Cardiff University, Cardiff, UK; and [2]School of Earth, Ocean and Planetary Sciences, Cardiff University, Cardiff, UK

Correspondence: John C. Fry, Cardiff School of Biosciences, Cardiff University, Main Building, Park Place, Cardiff, CF10 3TL, UK. Tel.: +44 29 2087 4190; fax: +44 29 2087 4305; e-mail: fry@cardiff.ac.uk

Received 14 December 2005; revised 9 March 2006; accepted 11 March 2006.
First published online 10 May 2006.

DOI:10.1111/j.1574-6941.2006.00144.x

Editor: Jim Prosser

Keywords
prokaryotic biodiversity; community composition; geochemistry; deep biosphere; denaturing gradient gel electrophoresis profiles; statistical analysis.

Abstract

The aim of this work was to relate depth profiles of prokaryotic community composition with geochemical processes in the deep subseafloor biosphere at two shallow-water sites on the Peru Margin in the Pacific Ocean (ODP Leg 201, sites 1228 and 1229). Principal component analysis of denaturing gradient gel electrophoresis banding patterns of deep-sediment *Bacteria*, *Archaea*, *Euryarchaeota* and the novel candidate division JS1, followed by multiple regression, showed strong relationships with prokaryotic activity and geochemistry ($R^2 = 55$–100%). Further correlation analysis, at one site, between the principal components from the community composition profiles for *Bacteria* and 12 other variables quantitatively confirmed their relationship with activity and geochemistry, which had previously only been implied. Comparison with previously published cell counts enumerated by fluorescent *in situ* hybridization with rRNA-targeted probes confirmed that these denaturing gradient gel electrophoresis profiles described an active prokaryotic community.

Introduction

Denaturing gradient gel elecrophoresis (DGGE) has been used extensively to profile prokaryotic community composition over both time and space in soils and aquatic environments (Schäfer & Muyzer, 2001). It provides a quicker, less labour-intensive approach to comparing community composition in many different samples than sequencing of clone libraries. Although primarily used with bacterial communities by amplifying fragments from 16S rRNA genes (Muyzer & Smalla, 1998), DGGE has also been used to explore the diversity of *Archaea* (Hoj *et al.*, 2005) and specific, nondominant, functional groups of prokaryotes (Dar *et al.*, 2005; Webster *et al.*, 2005). Traditionally, relationships between DGGE profiles and environmental variables have been suggested by tracking the appearance and disappearance of bands and linking these to other changes occurring in the ecosystem (Riemann *et al.*, 2000). Recently, DGGE profiles have been analysed successfully using a variety of statistical approaches (for a review see Fromin *et al.*, 2002). These techniques have included cluster analysis and ordination methods such as principal component analysis (PCA; Van der Gucht *et al.*, 2001; Yang *et al.*, 2001), multidimensional scaling (MDS; Schäfer *et al.*, 2001; Abell & Bowman, 2005b), and canonical variate analysis (McCaig *et al.*, 2001). In addition, ordination techniques such as canonical correspondence analysis allowed Yang & Crowley (2000) to show that 40% of the variation of bacterial community composition in a barley rhizosphere microcosm was a result of the iron nutritional status of the plant.

The predominant molecular approach to studying prokaryotic diversity in deep marine subseafloor sediments has been the sequencing of clone libraries obtained after PCR amplification of extracted DNA with primers amplifying fragments of genes from *Bacteria* (Rochelle *et al.*, 1994; Kormas *et al.*, 2003), *Archaea* (Sørensen *et al.*, 2004), and in some cases specific functional groups such as methanogens (Marchesi *et al.*, 2001; Newberry *et al.*, 2004). The labour-intensive nature of processing these samples and the difficulty of obtaining PCR-amplifiable DNA from these low-biomass sediments have contributed to the limited use of DGGE approaches. However, recent developments allowing samples with a low DNA template to be readily amplified by nested PCR (Webster *et al.*, 2003) have enabled DGGE profiling of the community composition of prokaryotes with depth in some deep biosphere sediments. In addition,

we have recently reported a brief description using PCA of DGGE banding patterns, coupled with multiple regression and correlation, to relate community composition of *Bacteria* to prokaryotic activity, total cell numbers and geochemical variables in one deep subsurface sediment site (Parkes *et al.*, 2005). Furthermore, a companion paper to this report compares sequencing of DGGE bands and clone libraries in deep subseafloor sediments and concludes that DGGE gives a reasonable summary of community composition (Webster *et al.*, 2006). The aim of this paper is to investigate the application of statistical approaches to relating DGGE community composition profiles to prokaryotic activity and geochemistry in two geographically close, deep subseafloor sediment sites from the Peru Margin in the Pacific Ocean.

Materials and methods

Site description and sample handling

The sites and sampling methods used here have been described elsewhere (Shipboard Scientific Party, 2003; Webster *et al.*, 2006) and are only described briefly here. Deep sediments were taken from the Ocean Drilling Program Leg 201, microbiologically focused cruise with full contamination checks and aseptic procedures (Shipboard Scientific Party, 2003). Sites 1228 and 1229 in the Peru Margin, eastern Pacific Ocean, in 262 and 150.5 m of water, respectively, were studied. Whole round cores (WRC) were cut onboard ship from 6.75–187.3 m below the seafloor (mbsf) for site 1228 and from 6.7–157.2 mbsf for site 1229. For DGGE studies, the WRC were immediately frozen, transported and stored at −80 °C until processed. Samples for activity measurements were taken from adjacent WRC sections with sterile 5 mL syringes (luer-end removed) in a cold room (4 °C) with anaerobic and aseptic handling and sealed with sterile, black butyl-rubber Subaseals (William Freeman Ltd, Barnsley, UK). These were equilibrated anaerobically at 16 °C for approximately one day prior to further processing. Prokaryotic activity was measured by injecting the syringe mini-cores individually with ^{3}H-thymidine, ^{14}C-acetic acid and ^{14}C-bicarbonate for estimating heterotrophic growth, acetoclastic and CO_2-reduction-based methanogenesis, respectively (see following sections for details). Four syringe mini-cores were used at each depth for each isotope, with one frozen immediately after injection as a blank, and any activity was subtracted from the incubated samples. Isotope was injected laterally along the centre-line of the syringe using an injection rig (Parkes *et al.*, 1995) and then incubated under N_2 at 16 °C (close to the average downcore temperature) for various times, with longer incubations for deeper depths. Incubations were terminated and stored by freezing at −20 °C.

Thymidine incorporation into DNA

Samples were injected with 19 μL of [methyl-^{3}H]-thymidine (674 kBq, specific activity 3.18 TBq mmol^{-1}) for between 4 h and 2 days before freezing. Subsequent processing was by extruding syringe contents into 50 mL centrifuge tubes containing 10 mL of 10% (w/v) trichloroacetic acid (TCA) at 4 °C. The samples were centrifuged at 2000 ***g***, rinsed a further two times with 5% (w/v) TCA at 4 °C to remove unincorporated ^{3}H-thymidine, and then twice with 95% (v/v) ethanol at 4 °C to remove any tritiated lipid fraction. After drying the sample to remove residual ethanol, 7 mL of 1 M NaOH was added and the samples were incubated in a water bath at 37 °C for 1 h. Following centrifugation at 4 °C, the supernatant was separated and acidified [20% (w/v) TCA in 3.6 M HCl], and carrier DNA and RNA added (50 μL of a 1 mg mL^{-1} solution of herring sperm DNA or bovine RNA) together with a small amount of Kieselguhr. After further centrifugation at 4 °C, the supernatant, containing the ^{3}H-RNA fraction, was removed, and the remaining pellet rinsed twice with 5% (w/v) TCA at 4 °C. Finally, the pellet was incubated in a water bath at 100 °C in 5% (w/v) TCA to obtain the purified ^{3}H-DNA in solution. Radioactivity in a 5 mL subsample of supernatant was counted in OptiPhase HiSafe-3 scintillant (Perkin-Elmer, Beaconsfield, UK) placed in a Wallac 1414 WinSpectral liquid scintillation counter (Perkin-Elmer). Incorporation rates were directly calculated taking into account incubation times.

Rates of methanogenesis

Samples were injected with either 7.4 μL of [1,2-^{14}C]acetic acid, sodium salt (51 kBq, specific activity 2.11 Gbq mmol^{-1}) or 7.2 μL of sodium [^{14}C]bicarbonate (130 kBq, specific activity 2.07 Gbq mmol^{-1}) and incubated for 0.3–3 days (acetate) or 2–18 days (bicarbonate). Subsequently, samples were processed by extruding syringe contents into bottles containing 7 mL of 1 M NaOH and a magnetic stirring bar. These were immediately sealed with a butyl-rubber bung, allowed to defrost, and shaken so as to form a slurry. The headspace gas was sparged, while the slurry was stirred with a carrier gas (nitrogen/oxygen, 99 : 1) for 20 min through a CO_2 trap to ensure complete removal of any $^{14}CO_2$ carried over, and then over copper oxide in a furnace (Carbolite, Chelmsford, UK) at 800 °C. This oxidized $^{14}CH_4$ to $^{14}CO_2$, which was collected in two scintillation vials placed in a series containing OptiPhase HiSafe-3 scintillant with 7% (v/v) β-phenethylamine to absorb $^{14}CO_2$, and activity was measured in a scintillation counter. Potential rates of methanogenesis from bicarbonate were calculated from the label turnover time multiplied by the total CO_2 pool size (calculated from ODP Leg 201 pH and alkalinity results; Shipboard Scientific Party, 2003). Potential acetate methanogenesis calculations were similar,

except that pool size was for acetate (Shipboard Scientific Party, 2003) and disintegrations per minute (DPM) were doubled to account for the fact that the radiotracer was a 1 : 1 mixture of ^{14}C-methyl- and ^{14}C-carboxyl-labelled acetate.

DNA extraction, PCR amplification and DGGE methodology

Full details of methodology are provided elsewhere (Webster *et al.*, 2006). Briefly, DNA was extracted using a modified FastDNA Spin Kit for Soil (Qbiogene Inc., Carlsbad, CA) method and amplified by PCR aseptically using 16S rRNA gene primers for *Bacteria*, *Archaea*, *Euryarchaeota* and the candidate division JS1 (Webster *et al.*, 2004). Nested PCR was then used to provide small (<200 bp) DNA fragments for DGGE analysis using primers for *Bacteria* and *Archaea* as appropriate. These PCR products were separated using the DCode^{TM} Universal Mutation Detection System (Biorad Laboratories, Hercules, CA) on polyacrylamide gels with a urea/formamide gradient, run for 5 h to minimize run-time effects on band number and position (Sigler *et al.*, 2004).

Statistical analysis of DGGE banding patterns

DGGE bands were identified by visually inspecting gel images in Microsoft PowerPoint 2002 through a mask consisting of 33–48 horizontal slices, with each slice approximately the width of the brightest band. Brightness and contrast were adjusted for each image to facilitate band identification: a green-coloured, transparent mask proved easiest to use. DGGE bands were scored as present (score = 1) or absent (score = 0), and, although it has been suggested that qualitative, rather than quantitative, analysis of band density is preferable for DGGE analysis (Schäfer & Muyzer, 2001), we also scored bands for relative density using a 1–5 semiquantitative scale (1 = least dense, 5 = most dense, 0 = band absent). The data were analysed by PCA, MDS, factor analysis and cluster analysis using a variety of approaches. The data matrix for this analysis used depths as the variables, band scores as the values within each variable, and the correlation coefficient to calculate the similarity matrix. With general *Bacteria* primers, DNA from some depths could not be amplified without *Escherichia coli* contamination (Rochelle *et al.*, 1992; Kormas *et al.*, 2003), and all bands at the *E. coli* equivalent position were excluded from the analysis. All statistical analysis was done with Minitab Release 14 (Minitab Inc., State College, PA), except for MDS for which Primer 5 for Windows (http://www.pml.ac.uk/primer/primer5.htm) was used.

To interpret the principal components calculated from DGGE profiles the first three components were used as the dependent *y*-variable, and other prokaryotic and geochemical variables were used as the predictors (*x*-variables) in multiple regression analysis (Iles, 1993). First, stepwise multiple regression was used in which the *F*-statistic was set to 1.0 for both variable addition and removal. The stepwise solution was checked with best-subsets regression. In cases where multicollinearity was identified, separate regressions with the highly correlated variables used separately were undertaken. With this mix of approaches the best multiple regression solutions were identified. The unadjusted R^2 value (%) was used as the measure of variability explained by the regressions. Correlation analysis was used to identify further the main factors that might be affecting the DGGE depth profiles by correlating the first three principal components with the variables used in the multiple regression analysis.

Results

Prokaryotic activity and geochemistry of sites 1228 and 1229

Sites 1228 and 1229 are relatively close (~25 km apart), located in shallow water and organically rich (2–11%) (Patience *et al.*, 1990). In addition, they both have deep incursions of sulphate-rich brine resulting in increases in chloride (Shipboard Scientific Party, 2003) and sulphate (Fig. 1) in deeper layers to values well above normal seawater concentrations. Both sites showed subsurface peaks in alkalinity (Fig. 1) and dissolved inorganic carbon (mainly CO_2; Shipboard Scientific Party, 2003), indicating subsurface bacterial heterotrophic activity. Sulphate was high at the surface and decreased with depth until about 30–40 mbsf as a result of sulphate reduction. Sulphate concentrations did not reach zero at site 1228, but did at site 1229, and methane concentrations were lower at site 1228 than at site 1229. However, at both sites the methane maxima corresponded with the sulphate minima, demonstrating the expected competition between these groups of terminal oxidizing bacteria for common, limiting substrates, with energetic advantage to sulphate-reducing bacteria (SRB) in the presence of sulphate (Jørgensen, 2000). Overall, these profiles suggest lower subsurface activity at site 1228 than at site 1229, and this was supported by direct rate measurements (Fig. 1). Acridine orange direct counts (AODCs) of prokaryotes were only available for site 1229 (Parkes *et al.*, 2005) and showed distinct peaks, to much higher values than normally expected in deep sub-seafloor sediments, at ~30 and ~90 mbsf (1.2 and 95×10^8 cells cm^{-3}, respectively), which corresponded to the two methane/sulphate interfaces (Fig. 1a; Parkes *et al.*, 2005). Pore-water manganese (Fig. 1) and iron concentrations (Shipboard Scientific Party, 2003) were low at both sites in upper sediments but increased to 5–10 μM below ~70 mbsf at site 1228 and ~120 mbsf at site 1229. In addition, at site 1229 slight elevations of iron (to

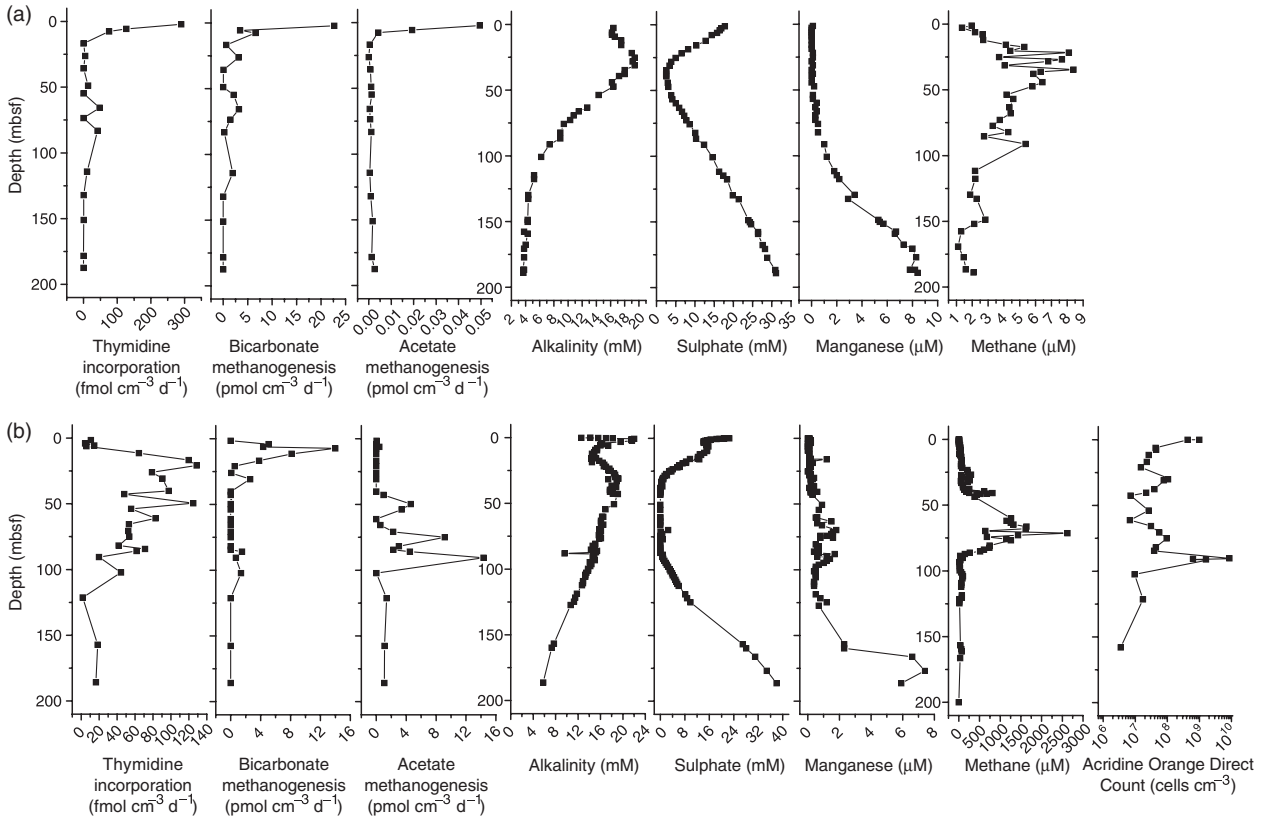

Fig. 1. (a) Activity and selected geochemical data for site 1228, and (b) selected geochemical data and AODCs for site 1229. Full details of the sites and scientific data are documented elsewhere (Shipboard Scientific Party, 2003), and the bacterial data for site 1229 have been described previously (Parkes *et al.*, 2005).

~14 μM) and manganese (to ~2 μM) concentrations were evident at ~70–90 mbsf. These pore-water dissolved metal concentration maxima suggest active bacterial metal reduction at considerable depth in both sites.

Here we report prokaryotic activity data for site 1228 for the first time. Data for site 1229 have been reported elsewhere (Parkes *et al.*, 2005) and are summarized here with additional geochemical data, so that they can be related to the DGGE analysis. At site 1228, thymidine incorporation (Fig. 1a) was high in near-surface sediments (290 fmol cm^{-3} day^{-1} at 2.26 mbsf) and decreased rapidly to zero at 16.7 mbsf. Below this depth, rates increased to 10–48 fmol cm^{-3} day^{-1} at depths of *c.* 50–115 mbsf, but again decreased to zero at greater depths. In contrast, thymidine incorporation at site 1229 (Fig. 1b; Parkes *et al.*, 2005) was low at the surface (~10 fmol cm^{-3} day^{-1}) but higher in subsurface layers. Maximum values of *c.* 90–120 fmol cm^{-3} day^{-1} occurred at 17–50 mbsf but then steadily decreased to ~50 fmol cm^{-3} day^{-1} until *c.* 100 mbsf, before declining to 2–18 pmol cm^{-3} day^{-1} at greater depths.

Bicarbonate methanogenesis was highest at site 1228 (Fig. 1a) in the top 8 m of sediment (3.5–23 pmol cm^{-3} day^{-1}), and below this was consistently low (mean = 1.4 pmol cm^{-3} day^{-1}) until 114 mbsf, below which it was undetectable. Acetate methanogenesis at site 1228 (Fig. 1a) followed a similar pattern, but rates were *c.* 1000-fold lower (0.05–0.0004 pmol cm^{-3} day^{-1}). Site 1229 was very different, with consistently higher subsurface rates (Fig. 1b; Parkes *et al.*, 2005) than site 1228, especially for acetate methanogenesis, and this was consistent with higher subsurface methane concentrations. Bicarbonate methanogenesis peaked at 7.45 mbsf (14.0 pmol cm^{-3} day^{-1}) whilst acetate methanogenesis, although undetectable between *c.* 7 and 40 mbsf, peaked at a much greater depth (14.3 pmol cm^{-3} day^{-1} at 90.51 mbsf). Despite low rates, acetate methanogenesis became the dominant methanogenic pathway below 100 mbsf (0.006–1.37 pmol cm^{-3} day^{-1}).

Comparison of ordination and clustering for DGGE band analysis

The DGGE profiles used for the analyses presented here are shown for site 1229 in Fig. 2 and for site 1228 in Fig. 3. Profiles are presented for the depths that gave amplification

products with the primer pairs used, and sequences from the marked bands have been identified and discussed in Webster *et al.* (2006).

Data from site 1229, obtained with general *Bacteria* primers, were used to check the efficacy of various statistical methods for analysis of DGGE banding patterns. A variety of ordination and clustering approaches demonstrated overall similarities between the results (Figs 4a–c). Generally, ordination methods showed more consistency than clustering approaches, with similar patterns using presence/absence and semiquantitative (data not shown) band scoring. Most of the cluster dendrograms gave a similar pattern to the single linkage plot using Manhattan distance (Fig. 4c). These ordination and clustering methods all identified profiles from 88.47 and 157.2 mbsf to be most different from the others. They also showed that the 6.7 and 11.2 mbsf profiles were distinctly different from those from the remaining depths, which clustered together to varying degrees. However, cluster analysis by Ward's method with squared Euclidean distance showed a very different branching order (Fig. 4d), failing to resolve the 88.47, 157.2, 6.7 and 11.2 mbsf profiles from the other depths.

These similarities show that the data are robust and that the analysis is consistent and reproducible. The ordination method using PCA with correlation from the presence/absence data matrix was chosen for all further work because of its simplicity and because the principal component loadings could then be analysed further with multiple regression and correlation. This allowed exploration of the links between prokaryotic community composition, cell numbers, activity and geochemistry.

PCA of DGGE profiles

PCA of all DGGE profiles from all depths gave good summaries of the data, as 54–79% of the variability was explained by the first three components (Table 1). Overall, more variability was explained by analyses for site 1229 (71–79%) than for site 1228 (54–67%). Two examples of component plots, both for site 1229, are given in Fig. 5. These are the DGGE profiles for which the plots were best explained by the data and show clear associations between sample depth and DGGE profile. For example, on both plots the shallowest depth profiles (6.7 and 11.2 mbsf) are distinct from the deepest profiles (157.2 mbsf in Fig. 5a and 120.7 mbsf in Fig. 5b) from which amplification products were obtained. In addition, all the profiles between 40.46 and 81.2 mbsf group together consistently. Furthermore, the depth profile closest to the AODC peak at 90 mbsf (88.47 mbsf; Fig. 1b) is separated from all other points for

Site 1229

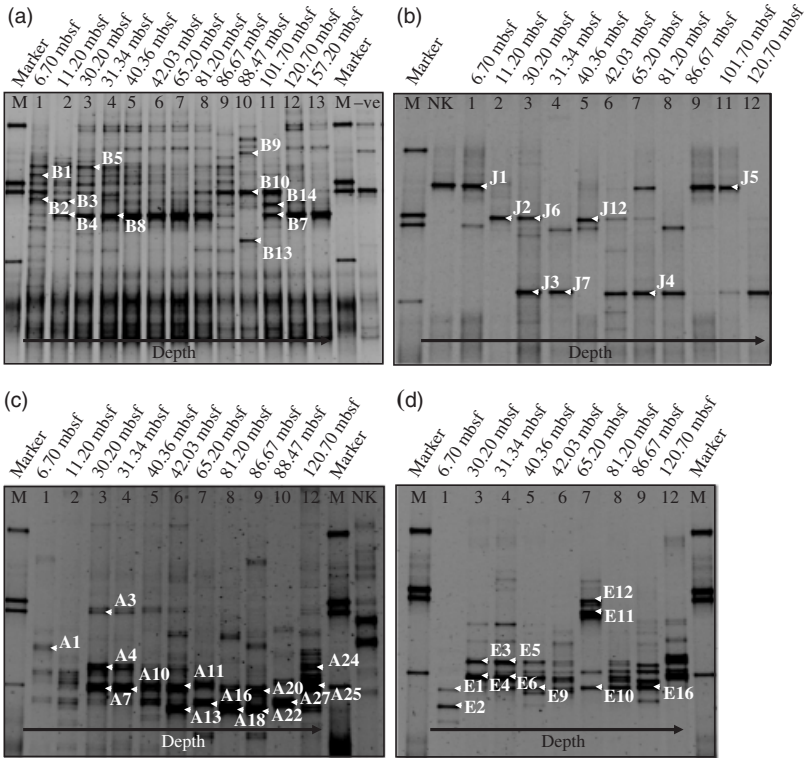

Fig. 2. DGGE analysis of 16S rRNA gene sequences from various sediment depths at the Peru Margin site 1229 by nested PCR. (a) Bacterial 16S rRNA genes using primers 27F-1492R and 357F-518R. (b) JS1 candidate division 16S rRNA genes using primers 63F-665R and 357F-518R. (c) Archaeal 16S rRNA genes using primers 109F-958R and SAf-PARCH519r. (d) Euryarchaeal 16S rRNA genes using primers 1A-1100A and SAf-PARCH519r. Lanes: M, DGGE marker (Webster *et al.*, 2003); -ve, nested PCR negative control (first-round PCR negative control re-amplified by nested PCR); NK, Nankai Trough sediment (4.15 mbsf) ODP Leg 190 site 1173 (Newberry *et al.*, 2004); 1 to 13, Peru Margin sediment site ODP Leg 201 site 1229 (lane numbers 1–13 are labelled above each panel by the depth of the profile). Marked DGGE bands were excised and sequenced (see Webster *et al.*, 2006 for band identities).

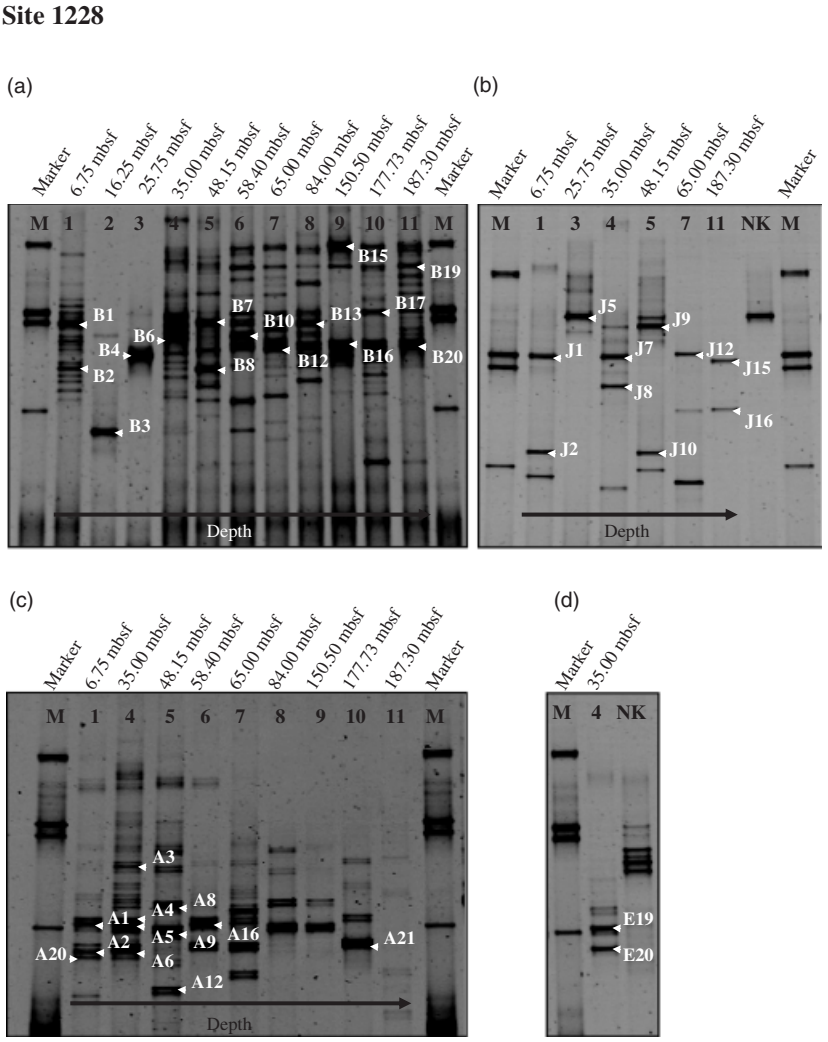

Fig. 3. DGGE analysis of bacterial and archaeal 16S rRNA gene sequences from various sediment depths at the Peru Margin site 1228 by nested PCR. (a) Bacterial 16S rRNA genes using primers 27F-1492R and 357F-518R. (b) JS1 candidate division 16S rRNA genes using primers 63F-665R and 357F-518R. (c) Archaeal 16S rRNA genes using primers 109F-958R and SAf-PARCH519r. (d) Euryarchaeal 16S rRNA genes using primers 1A-1100A and SAf-PARCH519r. Lanes: M, DGGE marker (Webster *et al.*, 2003); NK, Nankai Trough sediment (4.15 mbsf) ODP Leg 190 site 1173 (Newberry *et al.*, 2004); 1 to 13, Peru Margin sediment site ODP Leg 201 site 1229 (lane numbers 1–11 are labelled above each panel by the depth of the profile). Marked DGGE bands were excised and sequenced (see Webster *et al.*, 2006 for band identities).

the DGGE profiles of *Bacteria* (Fig. 5a), and the depths closest to the 30 mbsf AODC peak (30.20 and 31.34 mbsf) are close together on both plots.

Multiple-regression analysis of principal components and other variables

To interpret the DGGE banding patterns further in terms of other key variables, multiple regression was carried out between the first three components for all the PCAs using the prokaryotic and geochemical variables as predictors, together and separately (Table 1). In general, high proportions of the DGGE profile variability for PC1–3 could be explained by this approach. However, in many cases there were several equally good solutions, with different mixtures of variables accounting for similar proportions of variability. The components for *Bacteria* and *Archaea* were generally better explained for site 1229 (84–97%; mean = 91%) than for site 1228 (8–94%; mean = 61%); *Euryarchaeota* components were also well explained for site 1229 (82–93%). However, for the JS1 division the variabilities explained were higher for site 1228 (88–100%) than for site 1229 (47–87%). This poor explanation of the JS1 community composition profiles for site 1229 is consistent with what would be expected from the small number of band positions on the JS1 gel, namely 7 compared with 15–33 band positions on other gels (Table 1, Figs 2 and 3). Table 1 also shows that, when all variables were included in the regression analysis, in most cases a mixture of both prokaryotic and geochemical variables gave the best regression equations. It should be noted that AODCs were only conducted for site 1229, and that this variable occurs in only two of the 12 best equations using all variables for this site, and then only in the equation for PC3, which explains the least variability (Table 1). Overall, this analysis indicated that the DGGE profiles were related to both prokaryotic activity and geochemical variables. However, when these variables were used independently as predictors, in all cases except for the

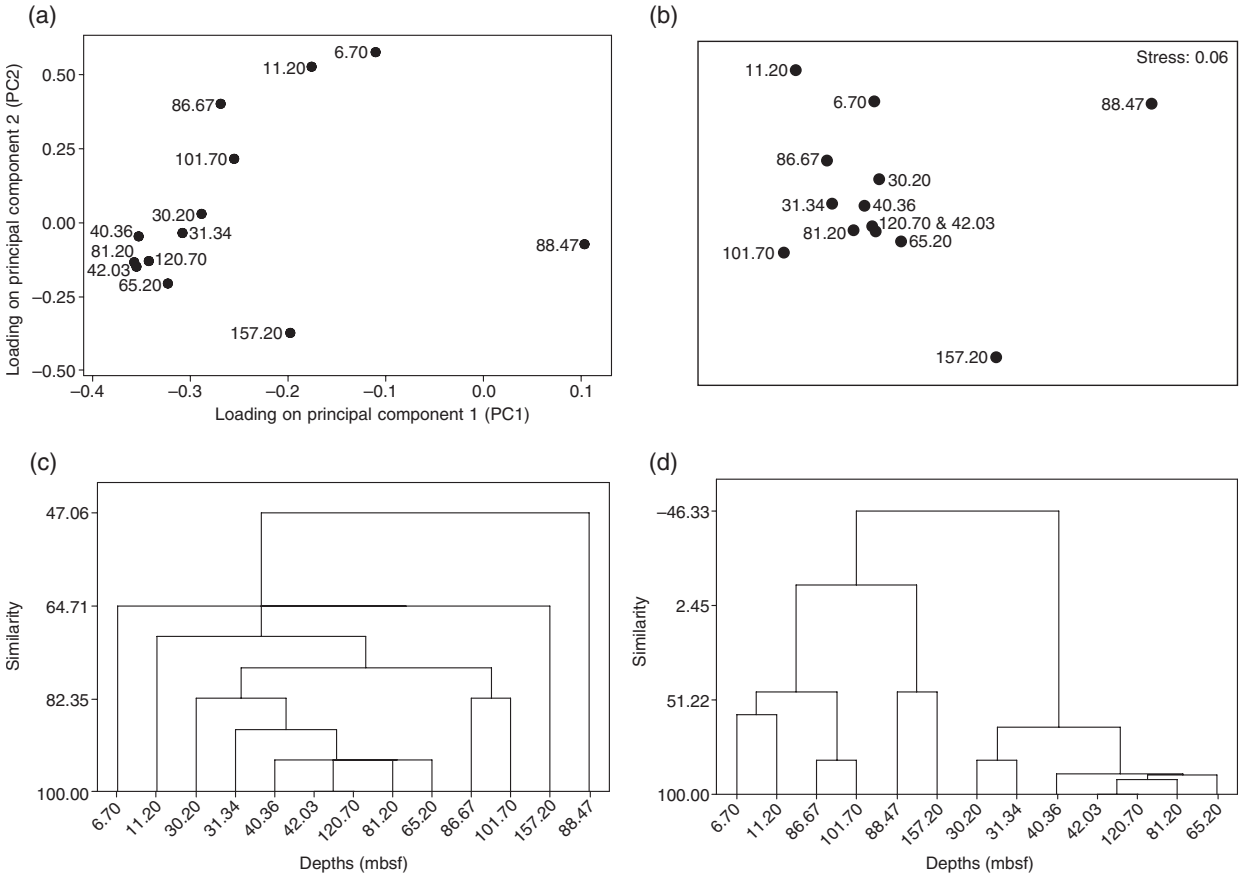

Fig. 4. Comparison of results from ordination and cluster analysis of DGGE banding profiles for *Bacteria* from site 1229, scoring bands as present or absent. The plots are (a) the loadings on the first two components from principal component analysis and (b) from multidimensional scaling, and (c,d) dendrograms from cluster analysis using single linkage with Manhattan distance and Ward's method with squared Euclidean distance, respectively.

site 1229 *Bacteria* profiles, the geochemical variables were more successful (mean variability explained = 74%) at predicting the principal components than the prokaryotic variables (mean = 45%). This may reflect the fact that activity measurements are relatively instantaneous measures of activity, whilst geochemical profiles arise from activity integrated over long time periods and that community composition is relatively stable over these timescales. Hence, the relationships are complex, and analysis suggests that several activity and geochemical variables interact with prokaryotic populations to define their community composition.

Correlation analysis of principal components with other variables

It was clear from the analysis of the *Bacteria* DGGE profiles from site 1229 (Figs 2a and 5a) that the 88.47 mbsf depth, closest to the 90 mbsf AODC peak (Fig. 1b), was extreme and so was likely to distort interpretations. So, to examine the underlying relationships for the community composition profiles across all depths, correlation analysis was undertaken between all the variables used in the regression analysis (Table 1) with PC1–3 from the DGGE profiles for *Bacteria* at site 1229, but with the data for the 88.47 mbsf depth excluded. The results (Table 2) confirm that AODCs are not strongly related to the overall community composition, as the correlations for this variable are not significant. In contrast, there were significant correlations with sulphate, barium, alkalinity, dissolved inorganic carbon (DIC), manganese, thymidine incorporation, bicarbonate and total methanogenesis (combined bicarbonate and acetate activities), and sulphate reduction rate (Parkes *et al.*, 2005; Table 2). Some of these variables were themselves correlated, for example sulphate and barium ($r = -0.727$, $P = 0.007$), alkalinity and DIC ($r = 0.949$, $P < 0.001$), and bicarbonate and total methanogenesis ($r = 0.834$, $P = 0.001$), as would be expected as all these pairs of variables are either geochemically or metabolically related. Thymidine incorporation is also significantly correlated with alkalinity and DIC ($r = 0.602$, $P = 0.038$ for alkalinity), which helps to confirm

Table 1. Results from multiple-regression analysis of components 1–3 (PC1–3) from principal-component analysis of the DGGE banding patterns of *Bacteria* and *Archaea* from sites 1229 and 1228

Component	Variability explained by component (%)	Variables* in best regression with all variables (variability explained†)	Variability explained (number of variables) in best regression with only	
			Prokaryotic variables	Geochemical variables
Bacteria, site 1229 (79.1%, 24‡)				
PC1	56.0	Tot-M (65%) Mn (76%) SRR (89%) Fe (92%)	72% (3)	71% (3)
PC2	13.7	Mn (35%) Tot-M (67%) SO_4 (78%) Cl (86%)	73% (3)	86% (4)
PC3	9.4	Alk (57%) Ac-M (73%) SRR (81%) Mn (84%)	78% (3)	65% (3)
Candidate division JS1, site 1229 (78.7%, 7)				
PC1	34.4	Fe (18%) S (47%)	53% (4)	68% (4)
PC2	24.5	S (28%) SRR (71%) TDR (84%) Cl (87%)	48% (4)	67% (4)
PC3	19.8	Bi-M (35%) CH_4 (47%)	39% (4)	78% (4)
Archaea, site 1229 (71.0%, 15)				
PC1	39.1	Tot-M (39%) CH_4 (55%) SO_4(91%) Mn (95%)	50% (3)	95% (4)
PC2	19.7	S (39%) Alk (70%) CH_4 (90%) Cl (92%)	31% (4)	90% (3)
PC3	12.2	Alk (35%) Fe (67%) CH_4(74%) SRR (84%) AODC (89%) Ac-M (97%)	53% (3)	82% (4)
Euryarchaeota, site 1229 (71.8%, 25)				
PC1	40.2	SO_4 (40%) SRR (59%) CH_4(77%) Bi-M (90%) TDR (93%)	88% (4)	90% (4)
PC2	18.0	SRR (24%) CH_4 (78%) SO_4 (86%) TDR (94%)	68% (4)	97% (4)
PC3	13.6	CH_4 (34%) Ac-M (53%) AODC (64%) Cl (82%)	70% (3)	63% (3)
Bacteria, site 1228 (54.4%, 32)				
PC1	23.9	NH_4 (19%) Bi-M (55%) Fe (73%)	16% (2)	59% (3)
PC2	17.4	TDR (23%) CH_4 (43%) Fe (64%) Ac-M (71%)	28% (2)	46% (4)
PC3	13.1	No significant regression possible	8% (2)	23% (3)
Candidate division JS1, site 1228 (66.8%, 19)				
PC1	25.9	CH_4 (33%) Ac-M (78%) TDR (88%)	35% (2)	86% (3)
PC2	21.8	S (39%) Fe (90%) Alk (96%) Bi-M (99%)	49% (2)	95% (3)
PC 3	19.1	Mn (67%) Cl (95%) Bi-M (100%) TDR (100%)	83% (2)	95% (2)
Archaea, site 1228 (62.6%, 33)				
PC1	29.6	SO_4 (42%) S (55%)	16% (2)	69% (4)
PC2	19.2	Bi-M (44%) Fe (62%) TDR (75%) CH_4 (80%) S (94%)	59% (2)	65% (4)
PC3	13.8	Fe (31%) SO_4 (70%)	11% (2)	70% (2)

*Variables used and their abbreviations in the tables were as follows.
Prokaryotic variables: thymidine incorporation rate (TDR), acetate methanogenesis (Ac-M), bicarbonate methanogenesis (Bi-M), total methanogenesis (Tot-M), sulphate reduction rate (SRR), $\log_{10}$ acridine orange direct count (AODC). Geochemical variables: methane (CH_4), alkalinity (Alk), chloride (Cl), sulphate (SO_4), ammonium (NH_4), iron (Fe), manganese (Mn), sulphide (S). All variables were used for site 1229, and all except AODC and SRR for site 1228, as these variables were not available. Values for all bacterial variables from site 1229 are published elsewhere (Parkes *et al.*, 2005) as are the geochemical variables (Shipboard Scientific Party, 2003).
†The variability explained is the total variability given sequentially for all the variables included in the best regression; further improvements by adding other variables were minimal.
‡These values are (i) the total percentage variability explained by the first three components estimated in the principal-component analysis, and (ii) the number of band positions showing at least one visible band at one depth interval that was used for the principal-component analysis.

that this variable predominantly measures productivity or growth associated with anaerobic heterotrophic activity, and not with sulphate reduction or methanogenesis (Wellsbury *et al.*, 1993). Thymidine incorporation was also negatively correlated with manganese concentration ($r = -0.729$, $P = 0.007$), and so was unlikely to be directly related to the growth of bacteria using manganese oxides as electron acceptors.

Only one significant correlation out of 108 was obtained between PC1–3 and the other variables in Table 2 for site 1229 archaeal, euryarcheotal and JS1 DGGE profiles, and this was between acetate methanogenesis and PC2 ($r = 0.922$, $P < 0.001$) from the archaeal profile analysis. There were, moreover, no significant correlations for *Bacteria*, *Archaea* or JS1 banding profiles for site 1228; however, for this site all the depths were used

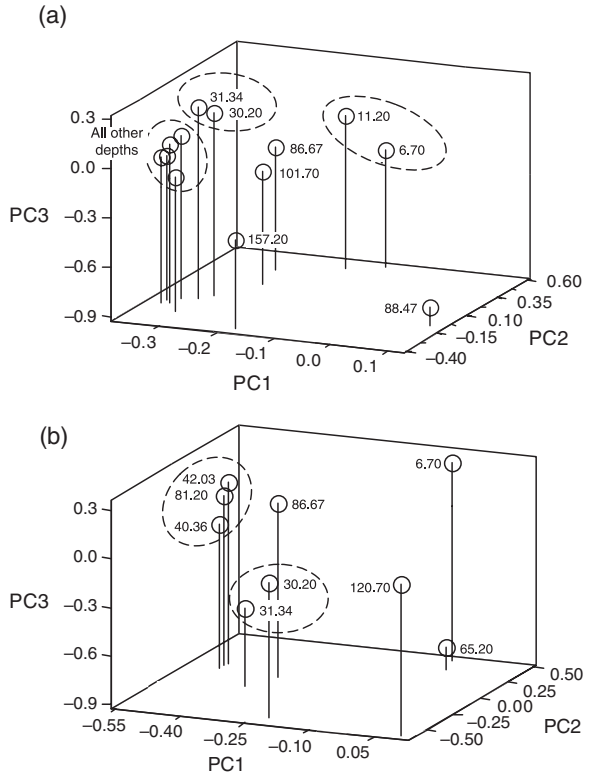

Fig. 5. Three-dimensional plots of the loadings for the first three principal components (PC1–3) from the analysis of DGGE gels of amplified DNA fragments extracted from site 1229 using primers for (a) *Bacteria* and (b) *Euryarchaeota*. The depths represented by the points plotted are indicated; dotted lines represent groups of points discussed in the text.

because there was no extreme community composition change at any depth that would justify exclusion from the analysis.

Discussion

This paper focuses on using multivariate analysis of banding patterns on DGGE gels to summarize, and so simplify, community composition data obtained using ordination and then on quantitatively relating the results to other prokaryotic and geochemical variables. It is therefore reasonable to discuss whether the amount of data generated by our DGGE gels was sufficient for this purpose. All but one (JS1, site 1229) of the gel profiles used showed large numbers of distinguishable bands (15–33 band positions, up to 18 bands per lane), which when related to other variables by PCA and multiple regression explained 55–97% (mean = 81%) of their variability. Furthermore, banding patterns within gels are known to be consistent and reproducible (Schauer *et al.*, 2000; Powell *et al.*, 2003), although reproducibility between gels is lower and therefore gels were analysed separately in our study (Powell *et al.*, 2003). In addition, the number of bands observed was similar to that found in other DGGE studies with 16S rRNA gene fragments (Schauer *et al.*, 2000; Powell *et al.*, 2003; Abell & Bowman, 2005a), where approximately 17–35 bands or positions per DGGE gel have been found. In our study there was no relationship between the number of band positions per gel and the percentage variability explained by the PCA. All these points together indicate that the DGGE analysis adopted was appropriate for the aims of this research.

As already stated, several other studies have used cluster analysis and ordination to summarize DGGE banding patterns. Most investigators score all the bands on a gel for their analysis; however, one study has selected the band positions to be investigated by ordination based on sequence identity (Hoj *et al.*, 2005). Similarly, most have used

Table 2. Correlation coefficients, and their significance†, between the first three principal components from DGGE band analysis of *Bacteria* 16S rRNA gene fragments amplified from site 1229 samples and selected geochemical and prokaryotic variables

Type of variable	Variable	PC1	PC2	PC3
Geochemical	Sulphate (mM)	**0.739****	0.057^{NS}	**−0.621***
	Barium (μM)	**−0.660***	−0.321^{NS}	0.265^{NS}
	Methane (μM)	−0.554^{NS}	−0.368^{NS}	0.197^{NS}
	Alkalinity (mM)	−0.291^{NS}	0.197^{NS}	**0.710****
	DIC (mM)	−0.505^{NS}	0.086^{NS}	**0.702***
	Manganese (μM)	0.052^{NS}	**−0.616***	**−0.604***
Prokaryotic	AODC (log_{10} cells cm^{-3})	−0.067^{NS}	0.448^{NS}	0.518^{NS}
	Thymidine incorporation (pmol cm^{-3} d^{-1})	−0.364^{NS}	−0.009^{NS}	**0.709****
	Acetate methanogenesis (pmol cm^{-3} d^{-1})	−0.168^{NS}	0.037^{NS}	−0.230^{NS}
	Bicarbonate methanogenesis (pmol cm^{-3} d^{-1})	**0.667***	**0.788****	0.167^{NS}
	Total methanogenesis (pmol cm^{-3} d^{-1})	**0.576***	**0.816*****	0.036^{NS}
	Sulfate reduction rate (pmol cm^{-3} d^{1})	**0.664***	0.543^{NS}	−0.242^{NS}

Note that all the values for the 88.47 mbsf depth were omitted ($n = 12$).

*,**,***: significance at $P = 0.05$, 0.01 and 0.001, respectively; significant coefficients are printed in bold.

DIC, dissolved inorganic carbon;

AODC, acridine orange direct count; NS, not significant.

presence and absence of bands at the positions identified to produce a binary matrix for analysis, but quantitative data from baseline-corrected gel scans are sometimes used (Yang & Crowley, 2000). Although it has been reported that analysis of a binary matrix is more reliable than the use of quantitative data (Schäfer & Muyzer, 2001), there appears to be no rigorous comparison published. Similarly, investigators usually select one method of analysis for their data and do not report comparative methods with different statistical analyses. This is why we compared several analytical approaches for the *Bacteria* profiles from site 1229, and it was reassuring to find that most pointed to a common overall result. Such consistency means that we can have confidence in the overall conclusions.

Despite the widespread use of multivariate approaches for summarizing DGGE profile data, few studies have used these approaches to express relationships between community composition and environmental variables as we have done. For example, using ordination Sekiguchi *et al.* (2002) reported temporal similarities at sites along the Changjiang River in China, and Hoj *et al.* (2005) concluded that methane emissions from arctic peat were more influenced by temperature and thaw depth than by archaeal community composition. Two other studies, both using plant/soil microcosms, related 40% of community composition to plant iron nutritional status (Yang & Crowley, 2000) and argued that heavy-metal- and plant-induced changes affected diversity (Gremion *et al.*, 2004). One study of benthic mat communities (Rothrock & Garcia-Pichel, 2005) showed that DGGE-based Shannon–Weaver diversity estimates of *Cyanobacteria*, *Bacteria* and *Archaea* showed strong negative correlations with tidal height ($R^2 = 0.535 - 0.967$).

The analyses that we present in the current paper go beyond the above studies and relate the DGGE profiles to a much wider range of environmental variables. For one DGGE profile (*Bacteria* at site 1229) we were able to relate the principal components from the PCA of the DGGE profile to other prokaryotic and geochemical variables (Table 2), and from this analysis Table 3 could be inferred. Some examples will demonstrate how this was possible. PC1 significantly and positively correlates with sulphate and sulphate reduction rate, and the only depth at which sulphate concentration and reduction rate are both high is in the near-surface layer (Fig. 1b; Parkes *et al.*, 2005). Similarly, bicarbonate and total methanogenesis correlate with PC1 and are both highest in the near subsurface (Fig. 1b; Parkes *et al.*, 2005). Therefore PC1 must be a near-surface/subsurface associated community composition variable. Although the correlation of PC1 with DIC and methane is not significant, their values are negative (−0.505 and −0.554, respectively), and their absolute values are higher than the other correlation coefficients. This supports the idea that carbon dioxide is the main substrate for methanogenesis at the surface and that methane is utilized in this zone (shown by high methanogenesis and very low methane concentration), perhaps by anaerobic methane oxidation (AMO; Parkes *et al.*, 2005). Using analogous logic, it appears that PC2 and PC3 have elements of activities at all the other depth zones identified in Table 3. For example, alkalinity, DIC and thymidine incorporation all correlate (close to 0.7) with PC3, which agrees with the hypothesis that heterotrophic activity and growth are important processes where thymidine incorporation is highest (about 15–90 mbsf). Manganese concentrations also correlated negatively with both PC2 and PC3, showing that, as these community composition components increase, manganese oxide utilization also increases, thus supporting the theory that metal reduction is an important prokaryotic activity in deeper layers (D'Hondt *et al.*, 2004; Parkes *et al.*, 2005).

Most of the sequences from the clone libraries and DGGE gels identify uncultured groups (Webster *et al.*, 2006), from which physiology cannot be inferred. However, the *Proteobacteria* sequences, which are well represented, are all phylogenetically close to known heterotrophs. This agrees with our conclusion for site 1229 that heterotrophy is important between *c.* 15 and 100 mbsf. Furthermore, heterotrophy is important near the surface, as alkalinity and DIC peak in this zone (Fig. 1b; Shipboard Scientific Party, 2003) and mesophilic metal-reducing bacteria are heterotrophs (Lovley *et al.*, 2004). Although *mcrA* genes were found at several depths (6.7, 30.2, 42.03 and 86.67 mbsf) they were not part of a dominant community (Parkes *et al.*, 2005; Webster *et al.*, 2006). Despite this, methanogenesis was a significant variable in many of our regression equations, implying that methanogenic activity affected community composition, perhaps indirectly.

It must be remembered that DGGE analysis of the type described here is only as reliable as the methodology. For example, whether the DNA comes from active or inactive cells, DNA extraction efficiency, primer selection, and other PCR biases all contribute to the banding pattern and relationships observed. Moreover, the observed bands at best represent the major prokaryotic groups targetted by the primers used, and minor groups with important geochemical influence may not be recovered. In fact, this has been observed for site 1229, as demonstrated by the near absence of methanogen and SRB sequences (Parkes *et al.*, 2005; Webster *et al.*, 2006) from this site, even though methanogenesis and sulphate reduction can clearly be inferred from geochemical profiles (D'Hondt *et al.*, 2004; Parkes *et al.*, 2005). The absence of AMO sequences in the clone libraries from site 1229 might be for similar reasons. Furthermore, we now know that the *Archaea* primers used for DGGE here were not ideal as they did not amplify the Miscellaneous Crenarchaeotic Group which dominated our

Table 3. Interpretation of depth zones in site 1229 from the analysis of the DGGE banding patterns of 16S rRNA genes of *Bacteria* in terms of prokaryotic activities based on bacterial and geochemical variables correlating (see Table 2) with principal components 1–3 (PC1–3)

Zone	State of variables in zone correlating with principal components from DGGE	Predominant activity(ies) deduced from DGGE analysis
Near-surface zone (top c.15 mbsf)	High sulphate reduction rate, low thymidine incorporation, highest sulphate, alkalinity and DIC, high bicarbonate methanogenesis	Sulphate reduction and bicarbonate methanogenesis
Upper deep active zone (c. 15–40 mbsf)	High thymidine incorporation, high alkalinity and DIC, decreasing sulphate and increasing Ba, low acetate methanogenesis and decreasing bicarbonate methanogenesis, low Mn	Heterotrophy*
Lower deep active zone (c. 40–100 mbsf)	Medium thymidine incorporation, decreasing alkalinity and DIC, low sulphate and high Ba, high acetate methanogenesis and low bicarbonate methanogenesis, elevated Mn	Acetate methanogenesis, heterotrophy* and metal reduction
Deep metal reduction zone (>100 mbsf)	Low thymidine incorporation, further decreasing alkalinity and DIC, rising sulphate and low Ba, low methanogenesis, rising Mn	Metal reduction

*Heterotrophy in this context refers to bacteria that are heterotrophic and can incorporate thymidine, so it excludes, for example, sulphate reducers. DIC, dissolved inorganic carbon; Mn, manganese; Ba, barium.

clone libraries, and were not found in any of the DGGE bands sequenced (Webster *et al.*, 2006).

Although the explanation summarized above and in Table 3 and the relationships identified between community composition, activity and geochemical variables are not perfect, the fact that many of the correlations described above for *Bacteria* from site 1229 are significant agrees with the overall message from regression analysis for all DGGE profiles from both sites and with conclusions in other papers (D'Hondt *et al.*, 2004; Parkes *et al.*, 2005). We have therefore shown that community composition is strongly related to prokaryotic activity and geochemistry in these sediments. This overall conclusion was consistent between two geographically close, but separate sites, and might be a general phenomenon within the deep marine biosphere. However, it is important to keep in mind that causality cannot be concluded from statistical analysis alone. In addition, the significant regression and correlation values that we have obtained might be the result of other, as yet unknown, variables being the real factors controlling community composition in the deep biosphere. Further studies are therefore required to examine the basis for the general relationships described, perhaps using new methodologies linking community composition and function (e.g. Madsen, 2005; Oremland *et al.*, 2005).

Other recent work using catalyzed reporter deposition FISH counting and quantitative PCR (Mauclaire *et al.*, 2004; Schippers *et al.*, 2005), and therefore targetting active organisms containing large numbers of ribosomes, shows clearly that both *Bacteria* and *Archaea* are important in the deep sub-seafloor biosphere. DGGE profiles therefore probably reflect active prokaryotes and not inactive remnants or fossil DNA from the past (DeLong, 2004; Parkes *et al.*, 2005).

Acknowledgements

This research used samples and data provided by the Ocean Drilling Program (ODP). Funding for this research was provided by the European Union DeepBUG project (contract number EVK3-CT-1999-017) and by the Natural Environment Research Council (NERC) Marine and Freshwater Microbial Biodiversity Programme (NER/T/S/2000/636). The authors would like to thank all shipboard scientists for their help in collecting Leg 201 samples, and Rachel Jones (Plymouth Marine Laboratory) for the MDS analysis of site 1229 DGGE data for *Bacteria*.

References

Abell GCJ & Bowman JP (2005a) Colonization and community dynamics of class Flavobacteria on diatom detritus in experimental mesocosms based on Southern Ocean seawater. *FEMS Microbiol Ecol* **53**: 379–391.

Abell GCJ & Bowman JP (2005b) Ecological and biogeographic relationships of class Flavobacteria in the Southern Ocean. *FEMS Microbiol Ecol* **51**: 265–277.

Dar SA, Kuenen JG & Muyzer G (2005) Nested PCR-denaturing gradient gel electrophoresis approach to determine the diversity of sulfate-reducing bacteria in complex microbial communities. *Appl Environ Microbiol* **71**: 2325–2330.

DeLong EF (2004) Microbial life breathes deep. *Science* **306**: 2198–2200.

D'Hondt S, Jørgensen BB, Miller DJ, Batzke A, Blake R, Cragg BA, *et al.* (2004) Distributions of microbial activities in deep subseafloor sediments. *Science* **306**: 2216–2221.

Fromin N, Hamelin J, Tarnawski S, Roesti D, Jourdain-Miserez K, Forestier N, Teyssier-Cuvelle S, Gillet F, Aragno M & Rossi P (2002) Statistical analysis of denaturing gel electrophoresis (DGE) fingerprinting patterns. *Environ Microbiol* **4**: 634–643.

Gremion F, Chatzinotas A, Kaufmann K, Sigler WV & Harms H (2004) Impacts of heavy metal contamination and phytoremediation on a microbial community during a twelve-month microcosm experiment. *FEMS Microbiol Ecol* **48**: 273–283.

Hoj L, Olsen RA & Torsvik VL (2005) Archaeal communities in high arctic wetlands at Spitsbergen, Norway (78 degrees N) as characterized by 16S rRNA gene fingerprinting. *FEMS Microbiol Ecol* **53**: 89–101.

Iles TC (1993) Multiple regression. *Biological Data Analysis* (Fry JC, ed), pp. 127–172. Oxford University Press, Oxford.

Jørgensen BB (2000) Bacteria and marine biogeochemistry. *Marine Geochemistry* (Schultz HD & Zabel M, eds), pp. 173–207. Springer-Verlag, Berlin.

Kormas KA, Smith DC, Edgcomb V & Teske A (2003) Molecular analysis of deep subsurface microbial communities in Nankai Trough sediments (ODP Leg 190, Site 1176). *FEMS Microbiol Ecol* **45**: 115–125.

Lovley DR, Holmes DE & Nevin KP (2004) Dissimilatory Fe(III) and Mn(IV) reduction. *Adv Microb Physiol* **49**: 219–286.

Madsen EL (2005) Identifying microorganisms responsible for ecologically significant biogeochemical processes. *Nat Rev Microbiol* **3**: 439–446.

Marchesi JR, Weightman AJ, Cragg BA, Parkes RJ & Fry JC (2001) Methanogen and bacterial diversity and distribution in deep gas hydrate sediments from the Cascadia Margin as revealed by 16S rRNA molecular analysis. *FEMS Microbiol Ecol* **34**: 221–228.

Mauclaire L, Zepp K, Meister P & Mckenzie J (2004) Direct *in situ* detection of cells in deep-sea sediment cores from the Peru Margin (ODP Leg 201, Site 1229). *Geobiology* **2**: 217–223.

McCaig AE, Glover LA & Prosser JI (2001) Numerical analysis of grassland bacterial community structure under different land management regimens by using 16S ribosomal DNA sequence data and denaturing gradient gel electrophoresis banding patterns. *Appl Environ Microbiol* **67**: 4554–4559.

Muyzer G & Smalla K (1998) Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie Van Leeuwenhoek Int J Gen Mol Microbiol* **73**: 127–141.

Newberry CJ, Webster G, Cragg BA, Parkes RJ, Weightman AJ & Fry JC (2004) Diversity of prokaryotes and methanogenesis in deep subsurface sediments from the Nankai Trough, Ocean Drilling Program Leg 190. *Environ Microbiol* **6**: 274–287.

Oremland RS, Capone DG, Stolz JF & Fuhrman J (2005) Whither or wither geomicrobiology in the era of 'community metagenomics'. *Nat Rev Microbiol* **3**: 572–578.

Parkes RJ, Cragg BA, Bale SJ, Goodman K & Fry JC (1995) A combined ecological and physiological approach to studying sulfate reduction within deep marine sediment layers. *J Microbiol Methods* **23**: 235–249.

Parkes RJ, Webster G, Cragg BA, Weightman AJ, Newberry CJ, Ferdelman TG, Kallmeyer J, Jørgensen BB, Aiello IW & Fry JC (2005) Deep sub-seafloor prokaryotes stimulated at interfaces over geological time. *Nature* **436**: 390–394.

Patience RJ, Clayton CJ, Kearsley AT, Rowland SJ, Bishop AN, Rees AWG, Bibby KG & Hopper AC (1990) An integrated biochemical, geochemical and sedimentological study of organic diagenesis in sediments from Leg 112. *Proceedings of the Ocean Drilling Program: Scientific Results* **112**: 135–146.

Powell SM, Bowman JP, Snape I & Stark JS (2003) Microbial community variation in pristine and polluted nearshore Antarctic sediments. *FEMS Microbiol Ecol* **45**: 135–145.

Riemann L, Steward GF & Azam F (2000) Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. *Appl Environ Microbiol* **66**: 578–587.

Rochelle PA, Weightman AJ & Fry JC (1992) DNAse-I treatment of taq DNA-polymerase for complete PCR decontamination. *Biotechniques* **13**: 520.

Rochelle PA, Cragg BA, Fry JC, Parkes RJ & Weightman AJ (1994) Effect of sample handling on estimation of bacterial diversity in marine sediments by 16S ribosomal-RNA gene sequence analysis. *FEMS Microbiol Ecol* **15**: 215–225.

Rothrock MJ & Garcia-Pichel F (2005) Microbial diversity of benthic mats along a tidal desiccation gradient. *Environ Microbiol* **7**: 593–601.

Schäfer H & Muyzer G (2001) Denaturing gel electrophoresis in marine microbial ecology. *Marine Microbiology* (Paul JH, ed), pp. 425–468. Academic Press, San Diego.

Schäfer H, Bernard L, Courties C, Lebaron P, Servais P, Pukall R, Stackebrandt E, Trousselier M, Guindulain T, Vives-Rigo J & Muyzer G (2001) Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: changes in the genetic diversity of bacterial populations. *FEMS Microbiol Ecol* **34**: 243–253.

Schauer M, Massana R & Pedros-Alio C (2000) Spatial differences in bacterioplankton composition along the Catalan coast (NW Mediterranean) assessed by molecular fingerprinting. *FEMS Microbiol Ecol* **33**: 51–59.

Schippers A, Neretin LN, Kallmeyer J, Ferdelman TG, Cragg BA, Parkes RJ & Jørgensen BB (2005) Prokaryotic cells of the deep sub-seafloor biosphere identified as living bacteria. *Nature* **433**: 861–864.

Sekiguchi H, Watanabe M, Nakahara T, Xu BH & Uchiyama H (2002) Succession of bacterial community structure along the Changjiang River determined by denaturing gradient gel electrophoresis and clone library analysis. *Appl Environ Microbiol* **68**: 5142–5150.

Shipboard Scientific Party (2003) Leg 201 summary. *Proceedings of the Ocean Drilling Program, Initial Reports, 201* (D'Hondt SL, Jørgensen BB & Miller DJ, *et al.*), pp. 1–81. Ocean Drilling Program, College Station, Texas.

Sigler WV, Miniaci C & Zeyer J (2004) Electrophoresis time impacts the denaturing gradient gel electrophoresis-based assessment of bacterial community structure. *J Microbiol Methods* **57**: 17–22.

Sørensen KB, Lauer A & Teske A (2004) Archaeal phylotypes in a metal-rich and low-activity deep subsurface sediment of the Peru Basin, ODP Leg 201, Site 1231. *Geobiology* **2**: 151–161.

Van der Gucht K, Sabbe K, De Meester L, Vloemans N, Zwart G, Gillis M & Vyverman W (2001) Contrasting bacterioplankton community composition and seasonal dynamics in two neighbouring hypertrophic freshwater lakes. *Environ Microbiol* **3**: 680–690.

Webster G, Newberry CJ, Fry JC & Weightman AJ (2003) Assessment of bacterial community structure in the deep sub-seafloor biosphere by 16S rDNA-based techniques: a cautionary tale. *J Microbiol Methods* **55**: 155–164.

Webster G, Parkes RJ, Fry JC & Weightman AJ (2004) Widespread occurrence of a novel division of bacteria identified by 16S rRNA gene sequences originally found in deep marine sediments. *Appl Environ Microbiol* **70**: 5708–5713.

Webster G, Embley TM, Freitag TE, Smith Z & Prosser JI (2005) Links between ammonia oxidizer species composition, functional diversity and nitrification kinetics in grassland soils. *Environ Microbiol* **7**: 676–684.

Webster G, Parkes RJ, Cragg BA, Newberry CJ, Weightman AJ & Fry JC (2006) Prokaryotic community composition in deep sub-seafloor sediments from the Peru Margin (ODP Leg 201, sites 1228 and 1229). *FEMS Microbiol Ecol* **58**: 65–85.

Wellsbury P, Herbert RA & Parkes RJ (1993) Incorporation of methyl-H-3 thymidine by obligate and facultative anaerobic-bacteria when grown under defined culture conditions. *FEMS Microbiol Ecol* **12**: 87–95.

Yang CH & Crowley DE (2000) Rhizosphere microbial community structure in relation to root location and plant iron nutritional status. *Appl Environ Microbiol* **66**: 345–351.

Yang CH, Crowley DE & Menge JA (2001) 16SrDNA fingerprinting of rhizosphere bacterial communities associated with healthy and Phytophthora infected avocado roots. *FEMS Microbiol Ecol* **35**: 129–136.